Generative AI for Health Economic Evaluation

IHEA | Robert Smith, PhD | October 2024

rsmith@darkpeakanalytics.com

https://github.com/dark-peak-analytics

https://www.linkedin.com/company/dark-peak-analytics

Limited peer reviewed literature

arXiv:2407.11054 [cs.LG] Sept 21, 2024 Version 3

Generative AI for Health Technology Assessment: Opportunities, Challenges, and Policy Considerations

Rachael L. Fleurence, PhD, MSc[1], Jiang Bian, PhD[2,3,4], Xiaoyan Wang, PhD[5,6], Hua Xu, PhD[7], Dalia Dawoud, PhD[8,9], Mitch Higashi, PhD[10], Jagpreet Chhatwal, PhD[11,12]

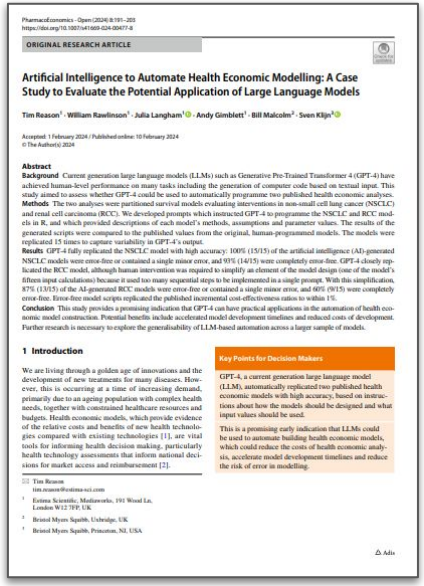
[1] Office of the Director, National Institute of Biomedical Imaging and Bioengineering, National Institutes of Health, Bethesda, MD, United States
[2] Health Outcomes and Biomedical Informatics, College of Medicine, University of Florida, Gainesville, FL, United [3] Biomedical Informatics, Clinical and Translational Science Institute, University of Florida, Gainesville, FL, United States
[4] Office of Data Science and Research Implementation, University of Florida Health, Gainesville, FL, United States
[5] Tulane University School of Public Health and Tropical Medicine, New Orleans, LA, United States
[6] Intelligent Medical Objects, Rosemont, IL, United States
[7] Department of Biomedical Informatics and Data Science, School of Medicine, Yale University, New Haven, CT, United States
[8] National Institute for Health and Care Excellence, London, United Kingdom
[9] Cairo University, Faculty of Pharmacy, Cairo, Egypt
[10] ISPOR - The Professional Society for Health Economics and Outcomes Research, Lawrenceville, NJ, United States
[11] Institute for Technology Assessment, Massachusetts General Hospital, Harvard Medical School, Boston, MA, United States
[12] Center for Health Decision Science, Harvard University, Boston, MA, United States

Funding: Dr Dalia Dawoud reports partial funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No. 82516 (Next Generation Health Technology Assessment (HTx) project. No other funding was received.

Acknowledgements: The authors thank Dr Tala Fakhouri for her comments on earlier versions of this manuscript.

1

Fleurence, R., Bian, J., Wang, X., Xu, H., Dawoud, D., Fakhouri, T., Higashi, M. and Chhatwal, J., 2024. Generative AI for Health Technology Assessment: Opportunities, Challenges, and Policy Considerations. arXiv preprint arXiv:2407.11054.



PharmacoEconomics - Open (2024) 8:191–203
https://doi.org/10.1007/s41669-024-00477-8

ORIGINAL RESEARCH ARTICLE

Artificial Intelligence to Automate Health Economic Modelling: A Case Study to Evaluate the Potential Application of Large Language Models

Tim Reason[1] · William Rawlinson[1] · Julia Langham[1] · Andy Gimblett[1] · Bill Malcolm[2] · Sven Klijn[3]

Accepted: 1 February 2024 / Published online: 10 February 2024

Abstract
Background Current generation large language models (LLMs) such as Generative Pre-Trained Transformer 4 (GPT-4) have achieved human-level performance on many tasks including the generation of computer code based on textual input. This study aimed to assess whether GPT-4 could be used to automatically programme two published health economic analyses.
Methods The two analyses were partitioned survival models evaluating interventions in non-small cell lung cancer (NSCLC) and renal cell carcinoma (RCC). We developed prompts which instructed GPT-4 to programme the NSCLC and RCC models in R, and which provided descriptions of each model's methods, assumptions and parameter values. The results of the generated scripts were compared to the published values from the original, human-programmed models. The models were replicated 15 times to capture variability in GPT-4's output.
Results GPT-4 fully replicated the NSCLC model with high accuracy: 100% (15/15) of the artificial intelligence (AI)-generated NSCLC models were error-free or contained a single minor error, and 93% (14/15) were completely error-free. GPT-4 closely replicated the RCC model, although human intervention was required to simplify an element of the model design (one of the model's fifteen input calculations) because it used too many sequential steps to be implemented in a single prompt. With this simplification, 87% (13/15) of the AI-generated RCC models were error-free or contained a single minor error, and 60% (9/15) were completely error-free. Error-free model scripts replicated the published incremental cost-effectiveness ratios to within 1%.
Conclusion This study provides a promising indication that GPT-4 can have practical applications in the automation of health economic model construction. Potential benefits include accelerated model development timelines and reduced costs of development. Further research is necessary to explore the generalisability of LLM-based automation across a larger sample of models.

1 Introduction

We are living through a golden age of innovations and the development of new treatments for many diseases. However, this is occurring at a time of increasing demand, primarily due to an ageing population with complex health needs, together with constrained healthcare resources and budgets. Health economic models, which provide evidence of the relative costs and benefits of new health technologies compared with existing technologies [1], are vital tools for informing health decision making, particularly health technology assessments that inform national decisions for market access and reimbursement [2].

Key Points for Decision Makers

GPT-4, a current generation large language model (LLM), automatically replicated two published health economic models with high accuracy, based on instructions about how the models should be designed and what input values should be used.

This is a promising early indication that LLMs could be used to automate building health economic models, which could reduce the costs of health economic analysis, accelerate model development timelines and reduce the risk of error in modelling.

✉ Tim Reason
tim.reason@estima-sci.com
1 Estima Scientific, Mediaworks, 191 Wood Ln, London W12 7FP, UK
2 Bristol Myers Squibb, Uxbridge, UK
3 Bristol Myers Squibb, Princeton, NJ, USA

Adis

Reason, T., Rawlinson, W., Langham, J., Gimblett, A., Malcolm, B. and Klijn, S., 2024. Artificial Intelligence to Automate Health Economic Modelling: A Case Study to Evaluate the Potential Application of Large Language Models. *PharmacoEconomics-Open*, *8*(2), pp.191-203.

Wellcome Open Research
Wellcome Open Research 2016 - DRAFT ARTICLE (PRE-SUBMISSION)

assertHE: an R package to improve quality assurance of HTA models

Robert A Smith[1,2], Yevgeniy Samyshkin[3], Wael Mohammed[1,2], Felicity Lamrock[2], Tom Ward[3,4], Jack Smith[1], Alan Martin[3], Paul Schneider[1,2], Dawn Lee[5], Gianluca Baio[6], Howard Thom[7], Nathan Green[8], Marina Richardson[11], Mohamed El Alili[12,13], Xavier Pouwels[8], Calum Lewis[1,2], and Baris Deniz[9,10]

[1] Dark Peak Analytics, Sheffield, UK
[2] University of Sheffield, Sheffield, UK
[3] GSK, London, UK.
[4] University of Exeter Medical School, Exeter, EX1 2LU, UK
[5] PenTAG, University of Exeter Medical School, University of Exeter, St Luke's Campus, Exeter, EX1 2LU, UK
[6] Department of Statistical Science, University College London, London, UK
[7] Department of Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, England, UK
[8] Section of Health Technology and Services Research, Technical Medical Centre, Faculty of Behavioural, Management, and Social Sciences, University of Twente, Enschede, Overijssel, The Netherlands
[9] Mathematical Sciences Research Centre, Queen's University Belfast, UK
[10] Current affiliation: Aide Solutions LLC, 2810 N Church St PMB 824887, Wilmington, Delaware 19802-4447, US
[11] Institute for Clinical and Economic Review (ICER), Boston, MA.
[12] Department of Health Sciences, Faculty of Science, Vrije Universiteit Amsterdam, Amsterdam Public Health research Institute, Amsterdam, The Netherlands
[13] National Health Care Institute (ZIN), Diemen, The Netherlands

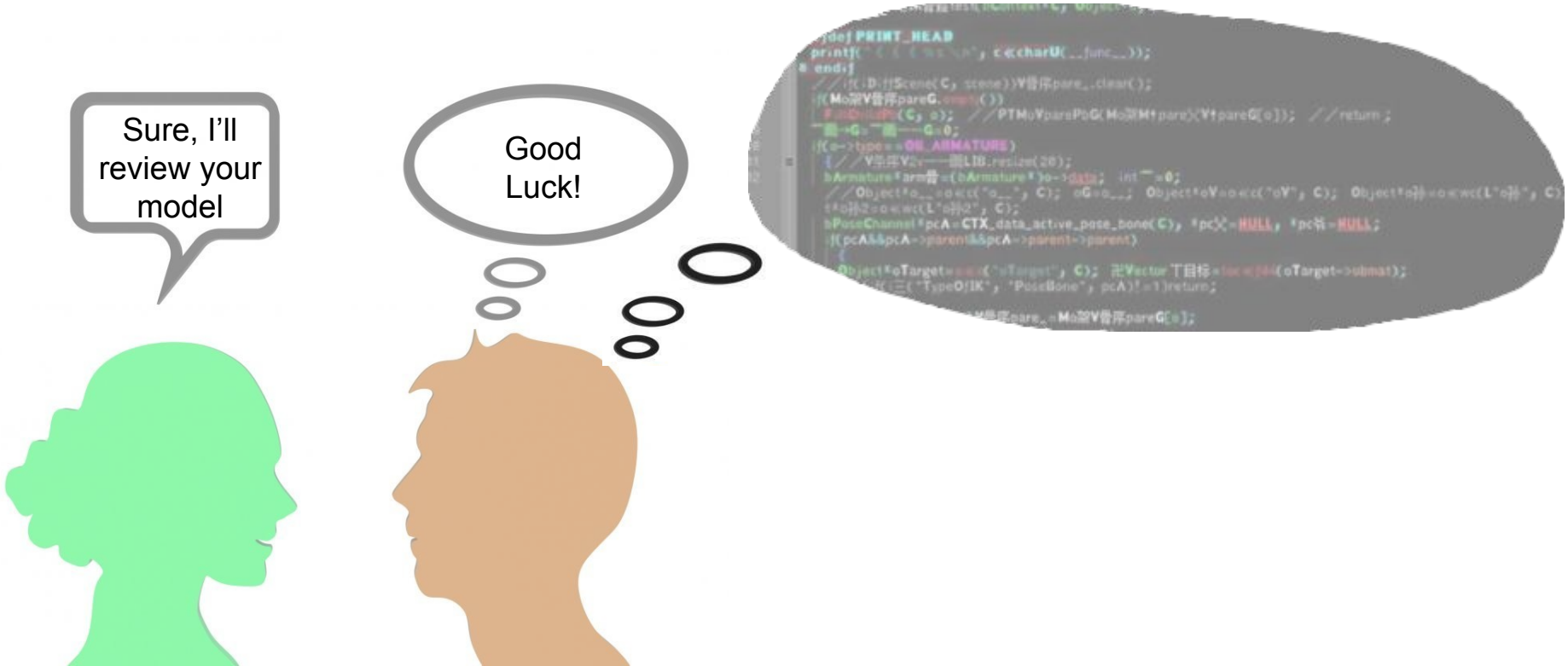
Corresponding author: Robert A Smith (rsmith@darkpeakanalytics.com).

Keywords
HTA, Health Economics, Unit Testing, Model Validation, Model Verification

Abstract
Background: Health economic models are increasingly used to inform decisions about the allocation of healthcare resources. Ensuring the robustness and reliability of these models is critical. Currently, quality assurance is conducted by both technical and non-technical experts assessing different components of the model manually. This is resource intensive. Understanding how the different components of the model fit together is time consuming, and testing every part of the model is sometimes not feasible in the time available. To aid in this, we have developed the assertHE R package.
Methods: The open source assertHE package provides testing functionality for those building and reviewing health economic models built in R programming language. It provides a series of checks which can be integrated into the model development workflow to reduce the probability of common errors. It also provides a suite of functions which allow users to better understand the network of R functions contained in a model, where they are defined, if (and where) they are tested, and provides a simple metric to quantify the extent to which they are tested.
Results: We applied the assertHE package to three open source health economic models built in R, showing how to include check functions within the model code and how to visualise the network of functions, see the test coverage, and obtain a Generative Pretrained Transformer Large Language Model (GPT-LLM) generated summary of any function in the code-base. We have worked with collaborators from industry, regulators and academia to develop the package to be applicable to the widest possible range of models, making adaptations to the source code based upon feedback.
Conclusions: assertHE offers an open-source toolkit for health economists building and reviewing models, promoting collaborative development and facilitating a more robust and efficient quality assurance process.

https://www.courses.darkpeakanalytics.com/products/digital_downloads/assertHE-quality-assurance

Background & Motivation


Sure, I'll review your model
Good Luck!

```r
year <- 0

#initialise the results matrix
results <- GenerateResultsMatrix(GlobalVars_, endtime_)
#initialise a PSA results matrix, if it is a PSA
if(GlobalVars_["run_psa","Value"]==T){
psaresults <- matrix(data=NA, nrow = as.numeric(GlobalVars_["Number of c
```

```r
calculate_costs <- function(population_, parameters_, year_, alive_, GlobalVars_, treatment_, attend_se_)

  population_[,"DMCOST"][alive_] <- population_[,"MET"][alive_]*parameters_[,"COST_MET"] +
      population_[,"MET2"][alive_] * parameters_[,"COST_MET2"] +
      population_[,"INSU"][alive_] * parameters_[,"COST_INSU"]

  population_[,"CVDCOST"][alive_] <-
      population_[,"IHD_E"][alive_] * parameters_[,"COST_IHD_E"] +
      population_[,"IHD_H"][alive_] * parameters_[,"COST_IHD_H"] +
      population_[,"MI_E"][alive_] * parameters_[,"COST_MI_E"] +
      population_[,"MI_H"][alive_] * parameters_[,"COST_MI_H"] +
      population_[,"MI2_E"][alive_] * (parameters_[,"COST_MI_E"] +
      #only add on the additional costs of a 2nd
      population_[,"STRO_E"][alive_] * parameters_[,"COST_STRO_E"] +
      population_[,"STRO_H"][alive_] * parameters_[,"COST_STRO_H"] +
      population_[,"STRO2_E"][alive_] * (parameters_[,"COST_STRO_E"] +
      #only add on the additional costs of a 2nd
      population_[,"CHF_E"][alive_] * parameters_[,"COST_CHF_E"] +
      population_[,"ATFIB_H"][alive_]*0+# note
```

```r
population_[,"AGE"] <- floor(diag_diab_population_[, "CURR_AGE"])
population_[,"MEN"] <- replace(population_[, "MEN"], population_[, "AGE"] > 51 & population_[, "FEMALE"
population_[,"HBA"] <- round(diag_diab_population_[, "HbA1c"],1)
#This has been kept the same as in the SPHR diabetes model
#Record the BMI of the population
#Apply a logical constraint to the BMI so that it cannot be less than 5kg/m2
population_[, "BMI"] <- replace(diag_diab_population_[, "BMI"], population_[, "BMI"] < 5, 5)
#Record cholesterol (add units)
population_[, "HDL"] <- diag_diab_population_[, "HDL"]
#Record systolic blood pressure(add units)
population_[, "SBP"] <- diag_diab_population_[, "SBP"]
```

```r
calculate_QALYs <- function(population_, parameters_, year_, alive_, GlobalVars_) {
#Calculate multiplier to adjust for a population with T2D without any complications
#Formula to estimate baseline QALYs
#Declare mean BMI, age and proportion female from Hayes et al
Mean_BMI_Hayes <-28.4 #Mean BMI in the source of our baseline utilities
Mean_Age_Hayes <- 65.8
Mean_pFemale_Hayes <- 4729/(6401+4729)

Util_bl_mult <- parameters_[,"UTIL_BL"] /
  (0.9454933+0.0256466*(1-Mean_pFemale_Hayes)+
  -0.0002213*Mean_Age_Hayes+
  -0.0000294*(Mean_Age_Hayes^2))

#Calculate utility prior to adjusting for BMI and events
population_[,"EQ5D"][alive_] <- (0.9454933 +
  0.0256466*population_[,"MALE"][alive_]+
  -0.0002213 * population_[,"AGE"][alive_]
  -0.0000294 * (population_[,"AGE"][alive_]^2))*Util_bl_mult

#apply the BMI decrements to this
population_[,"EQ5D"][alive_] <- population_[,"EQ5D"][alive_] +
  parameters_[,"UTIL_BMI"]*(population_[,"BMI"][alive_] - Mean_BMI_Hayes)
```

```r
##'@param random_numbers_ is the array of random
##'@return pBC is the vector of the probabilities of breast cancer
Breast_cancer <- function(population_, parameter_, alive_){
  #Calculate the fitted value
  FV <- parameter_[,"CANB_mu"]+
  parameter_[,"CANB_bta_MEN"]*population_[,"MEN"][alive_]+
  parameter_[,"CANB_bta_BMI"]*population_[,"BMI"][alive_]+
  parameter_[,"CANB_bta_BMIMEN"]*(population_[,"MEN"][alive_]*population_[,"BMI"][alive_])
  #convert to probabilities
  pBC <- 1-exp(-exp(FV))
  #set breast cancer risk to 0 for men
  pBC <- ifelse(population_[,"FEMALE"]==0,0,pBC)
  #remove temporary variables generated in the function
  rm(FV)
  #return the probabilities
  return(pBC)
}
```

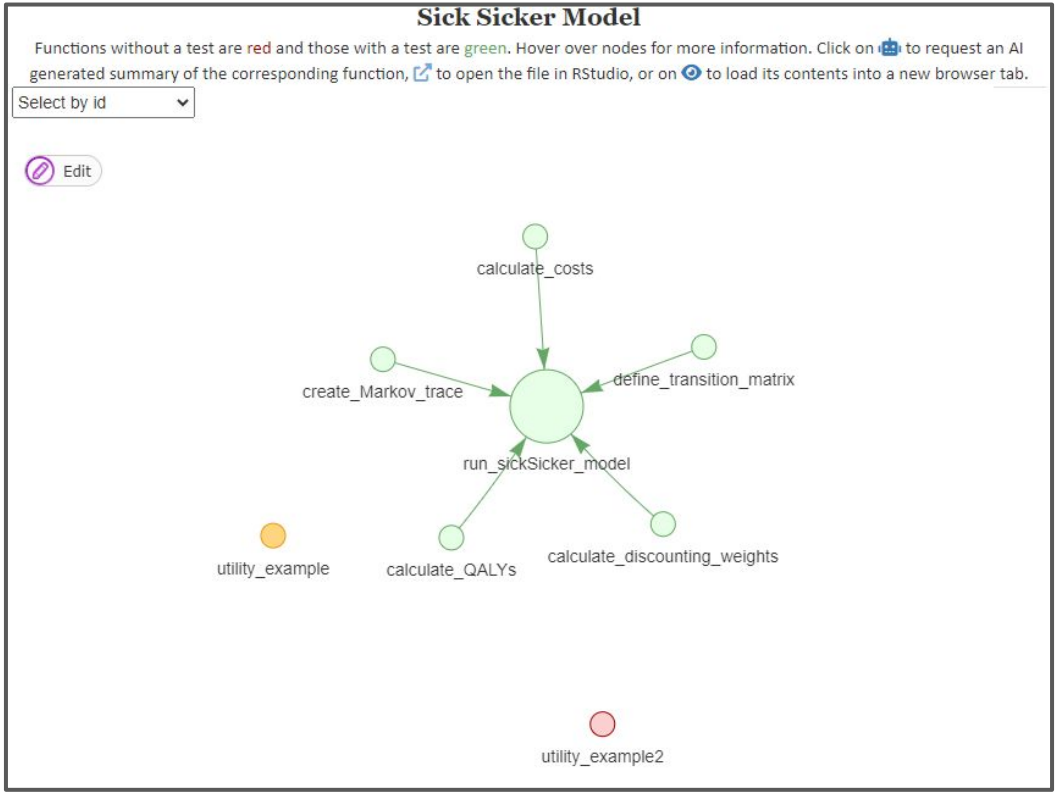
```r
#Set up the cluster
#to all the cluster
registerDoParallel(cl)
clusterExport(cl, ...)
modelresults <- parLapply(cl = cl,
                          PSAiterations,
                          run.simulation,
                          population_ = population,
                          parameters_ = parameters_,
                          endtime_ = endtime_,
                          treatment_ = treatment_,
                          GlobalVars_ = GlobalVars_,
                          random_numbs_ = random_numbers_,
                          LifeTables_ = LifeTables_)
stopCluster(cl)

if(GlobalVars["Results_output", "Value"] == "Summary"){
  modelresults <- matrix(unlist(modelresults), ncol=24, byrow=T)
```

- build_population.R
- Cancer Risks.R
- Costs.R
- Depression.R
- Generate Results Template.R
- Generate Results.R
- generate_random.R
- intervention.R
- LifeTableMortality.R
- Oestoarthritis functions.R
- QALYs.R
- Run model.R
- Run simulation.R
- UKPDS 82 risk functions.R
- UKPDS 90 risk functions.R
- Update Events.R
- Update Pat Chars.R

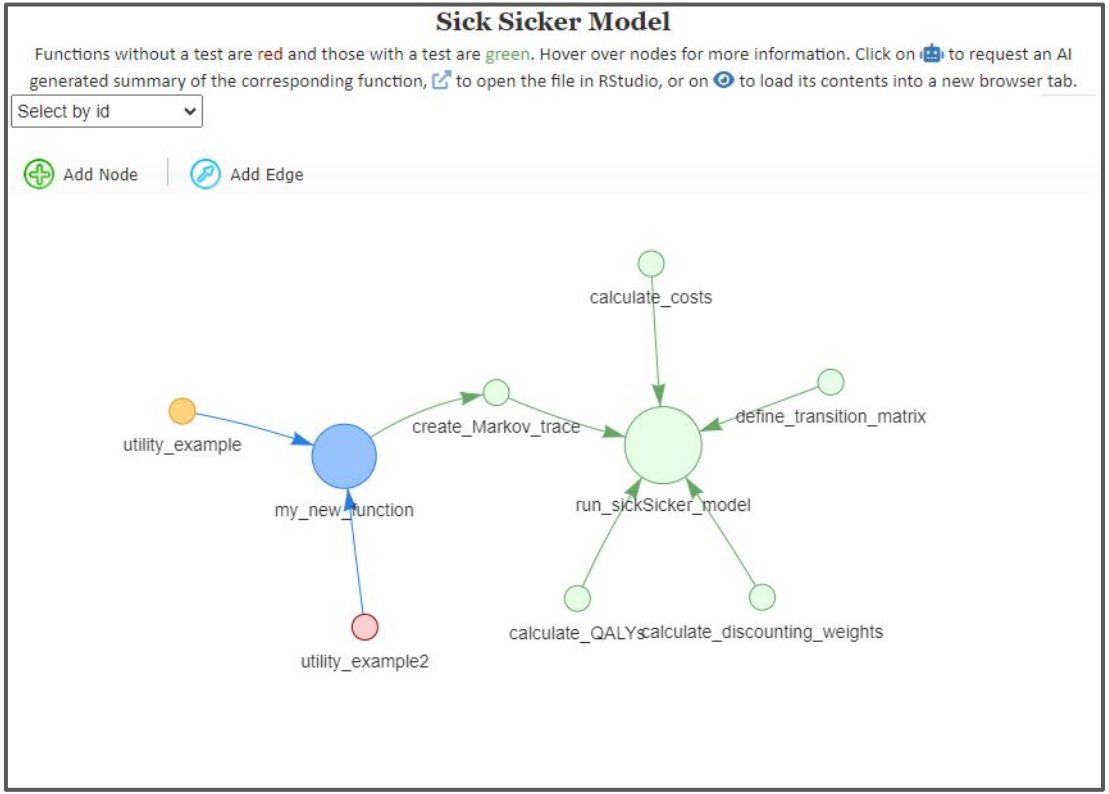
assertHE

assertHE model reviewer

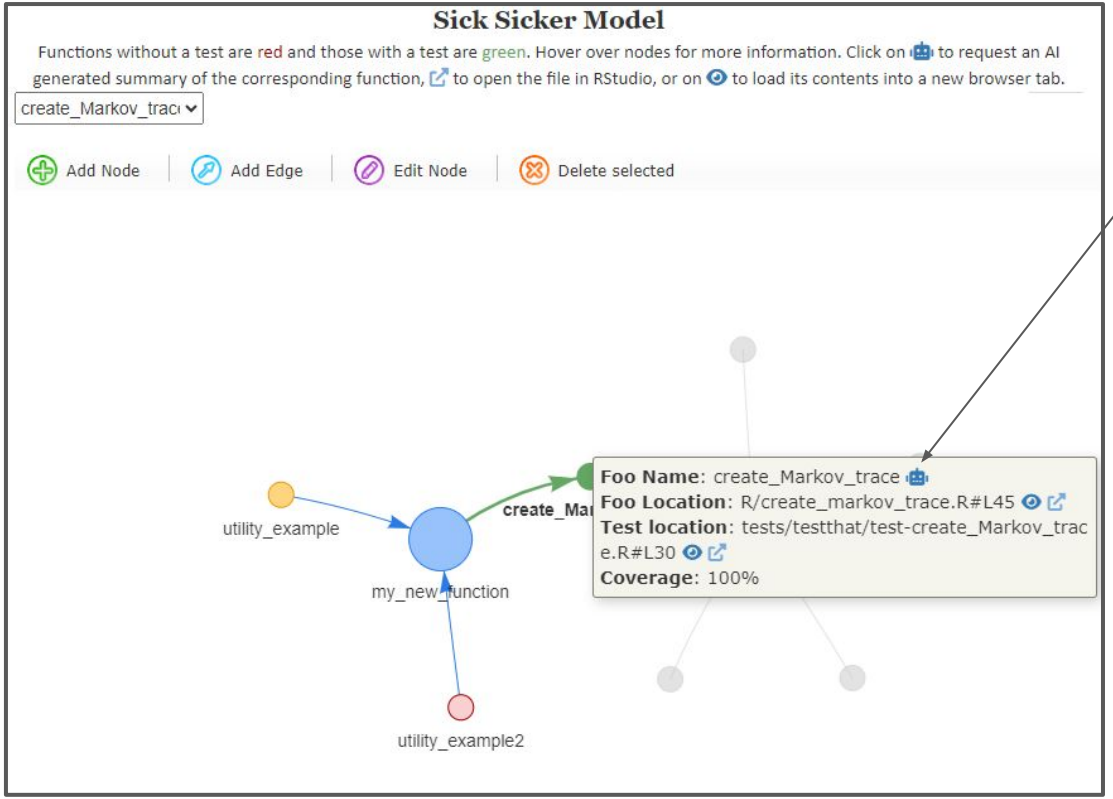


assertHE model reviewer

assertHE model reviewer

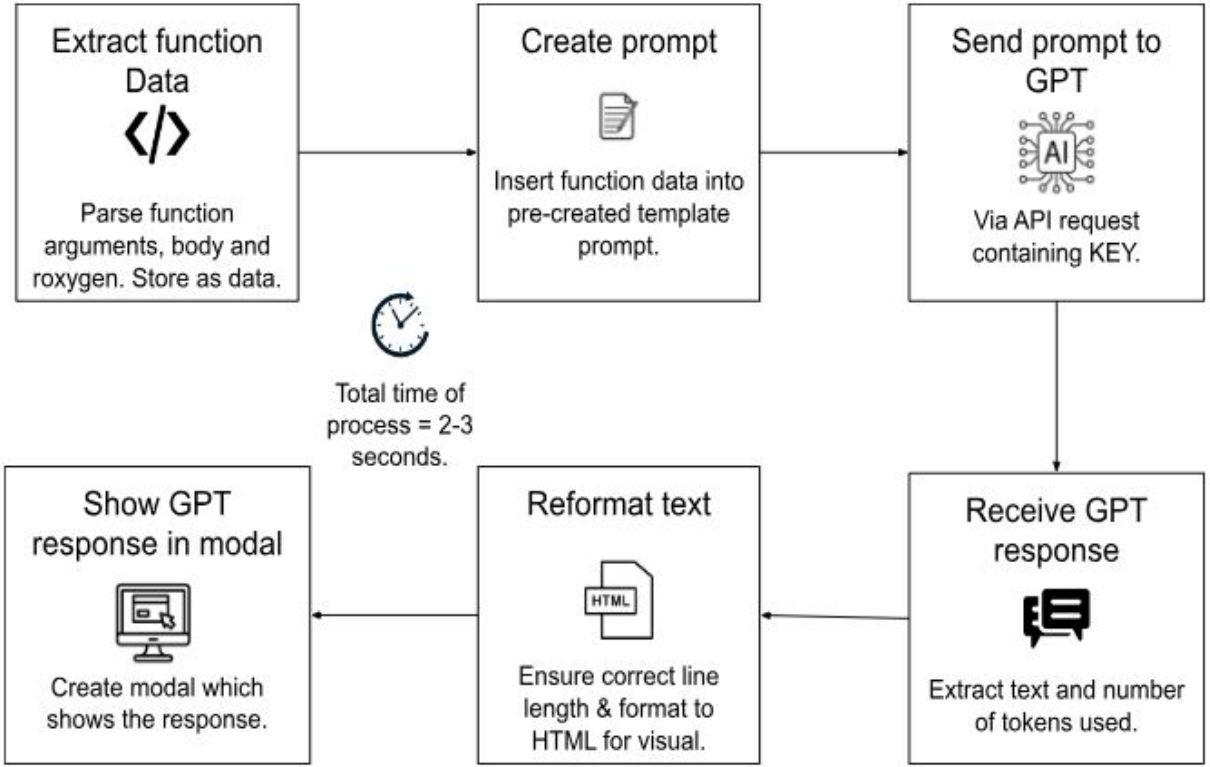


assertHE model reviewer

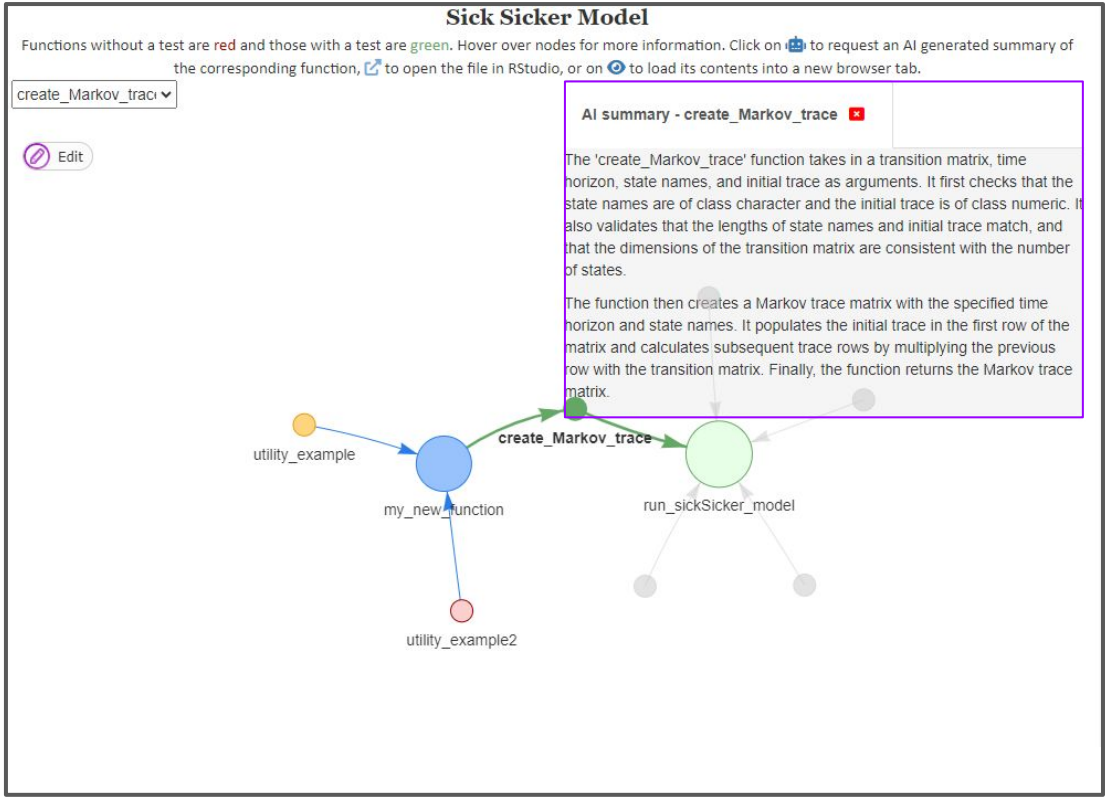


Generate LLM summary of function.

assertHE model reviewer



assertHE model reviewer



Case Studies

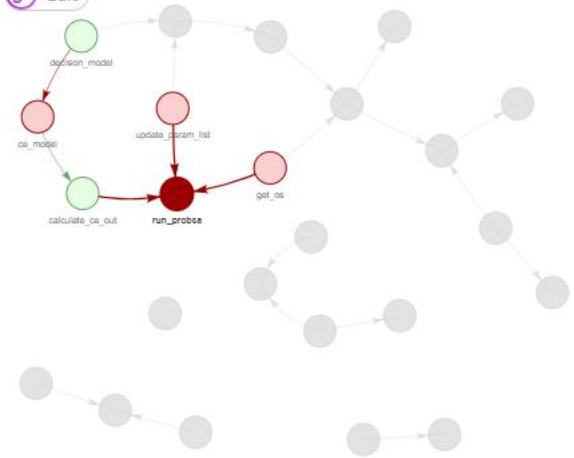
Case Study

Function Network

Functions without a test are red and those with a test are green. Hover over nodes for more information. Click on to request an AI generated summary of the corresponding function, to open the file in RStudio, or on to load its contents into a new browser tab.

run_probsa

Edit



AI summary - run_probsa

```html
The 'run_probsa' function runs probabilistic sensitivity analysis (PSA) on a given input dataset. If the 'parallel' argument is set to TRUE, the function parallelizes the PSA process using multiple cores based on the operating system. It then calculates costs and effects for each simulation, aggregates the results, and returns them in separate data frames.

If the 'parallel' argument is set to FALSE, the function runs the PSA simulations in series. It iterates through each simulation, updates parameters, calculates costs and effects, and prints the progress. Finally, it aggregates the results and returns them in separate data frames. The function returns a list containing the costs and effects data frames.
```

ELSEVIER

ScienceDirect

Contents lists available at sciencedirect.com

Journal homepage: www.elsevier.com/locate/jval

Economic Evaluation

CDX2 Biomarker Testing and Adjuvant Therapy for Stage II Colon Cancer: An Exploratory Cost-Effectiveness Analysis

Fernando Alarid-Escudero, PhD, Deborah Schrag, MD, MPH, Karen M. Kuntz, ScD

ABSTRACT

Objectives: Adjuvant chemotherapy is not recommended for patients with average-risk stage II (T3N0) colon cancer. Nevertheless, a subgroup of these patients who are CDX2-negative might benefit from adjuvant chemotherapy. We evaluated the cost-effectiveness of testing for the absence of CDX2 expression followed by adjuvant chemotherapy (fluorouracil combined with oxaliplatin [FOLFOX]) for patients with stage II colon cancer.

Methods: We developed a decision model to simulate a hypothetical cohort of 65-year-old patients with average-risk stage II colon cancer with 7.2% of these patients being CDX2-negative under 2 different interventions: (1) test for the absence of CDX2 expression followed by adjuvant chemotherapy for CDX2-negative patients and (2) no CDX2 testing and no adjuvant chemotherapy for any patient. We derived disease progression parameters, adjuvant chemotherapy effectiveness and utilities from published analyses, and cancer care costs from the Surveillance, Epidemiology, and End Results (SEER)–Medicare data. Sensitivity analyses were conducted.

Results: Testing for CDX2 followed by FOLFOX for CDX2-negative patients had an incremental cost-effectiveness ratio of $5500/quality-adjusted life-years (QALYs) compared with no CDX2 testing and no FOLFOX (6.874 vs 6.838 discounted QALYs and $89 991 vs $89 797 discounted US dollar lifetime costs). In sensitivity analyses, considering a cost-effectiveness threshold of $100 000/QALY, testing for CDX2 followed by FOLFOX on CDX2-negative patients remains cost-effective for hazard ratios of <0.975 of the effectiveness of FOLFOX in CDX2-negative patients in reducing the rate of developing a metastatic recurrence.

Conclusion: Testing tumors of patients with stage II colon cancer for CDX2 and administration of adjuvant treatment to the subgroup found CDX2-negative is a cost-effective and high-value management strategy across a broad range of plausible assumptions.

Keywords: CDX2, cost-effectiveness analysis, decision-analytic model, immunohistochemistry testing, stage II colon cancer.

VALUE HEALTH. 2022; 25(3):409–418

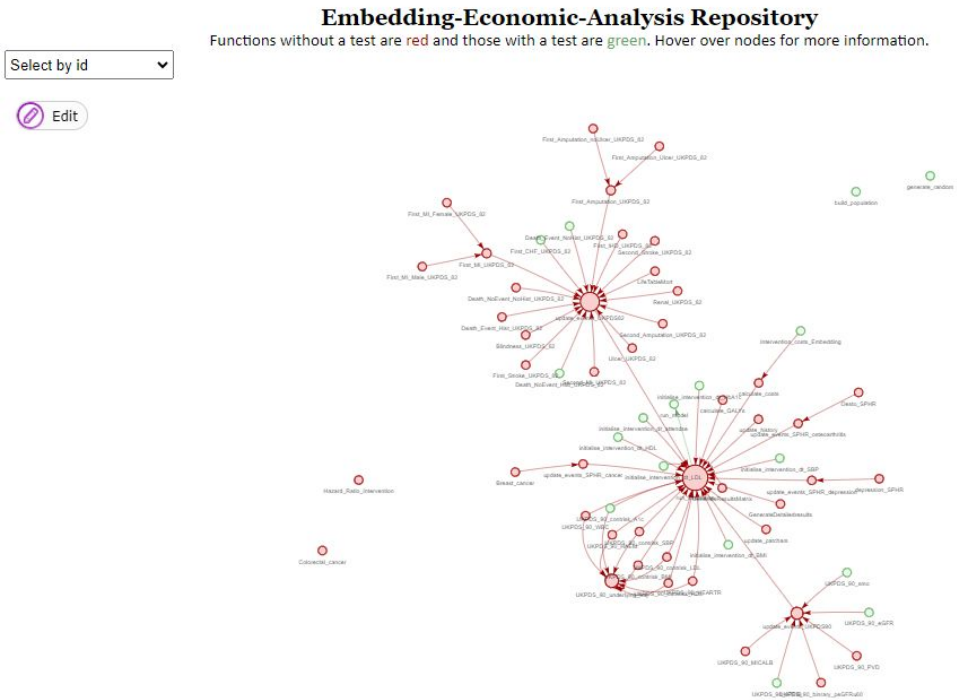
Introduction

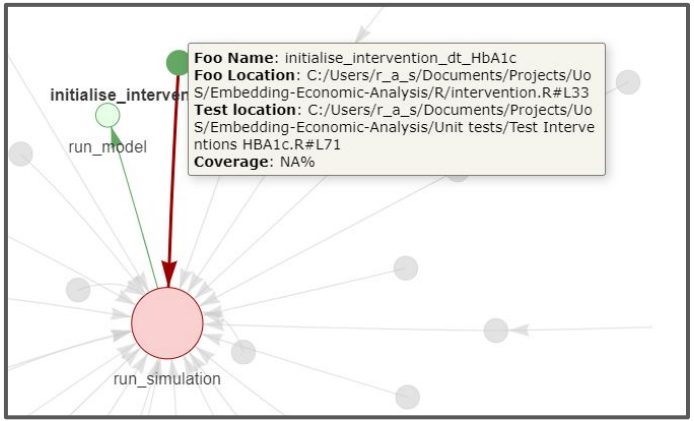
Adjuvant chemotherapy is not recommended for patients with average-risk stage II (T3N0) colon cancer,[1–4] and thus, these patients are usually treated with surgery alone.[5] Nevertheless, a recent study by Dalerba et al[6] described a small subgroup of patients with average-risk stage II colon cancer who lack expression of the CDX2 transcription factor that associated with clinical benefit from adjuvant chemotherapy. CDX2 is a master transcription factor involved in intestinal development[7] and serves as a candidate biomarker of mature colon epithelial tissues. In this study, the authors used Boolean implication networks[8] to conduct a systematic search for a biomarker to identify undifferentiated tumors in a collection of human colon gene expression array experiments from the National Center for Biotechnology Information Gene Expression Omnibus repository (www.ncbi.nlm.nih.gov/geo) used as a discovery data set. As a validation data set, the authors used tissue microarrays from the Cancer Diagnosis Program of the National Cancer Institute, which were analyzed for CDX2 expression by immunohistochemical (IHC) analysis. Among all tumors analyzed in the validation data set, 48 of 669 (7.2%) were CDX2-negative, defined as completely lacking CDX2 expression or showing expression in a minority of malignant epithelial cells.[6,9] The study also showed that CDX2-negative patients had poorer 5-year disease-free survival (DFS) than CDX2-positive patients (those with biomarker expression). More importantly, the 5-year DFS was greater for the CDX2-negative patients who received adjuvant chemotherapy than similar patients who did not receive adjuvant chemotherapy. The ability to test average-risk patients with stage II colon cancer for CDX2 biomarker expression to target adjuvant chemotherapy to a subgroup most likely to benefit could reduce colon cancer mortality and minimize adjuvant chemotherapy harms.[10]

This study aims to quantify the long-term benefits, costs, and cost-effectiveness of testing average-risk patients with stage II colon cancer for the absence of CDX2 biomarker expression followed

https://github.com/feralaes/cdx2cea

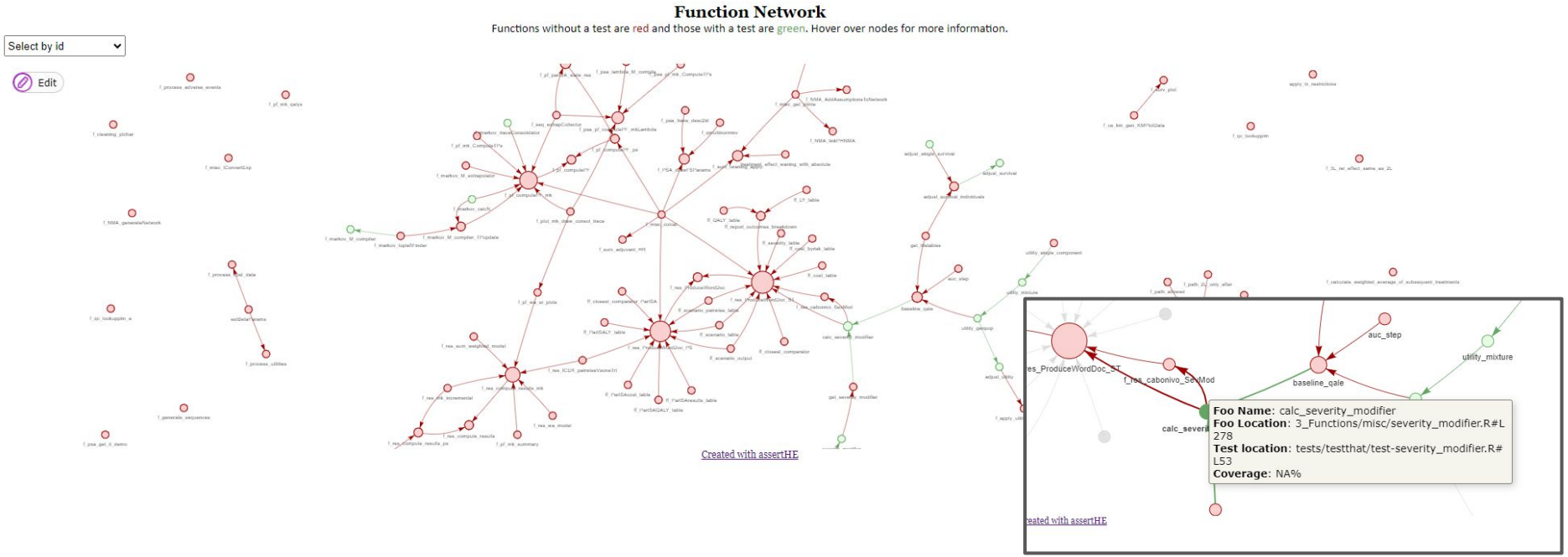
Case Study: Embedding Economic Analysis





https://github.com/DanPollardSheff/Embedding-Economic-Analysis

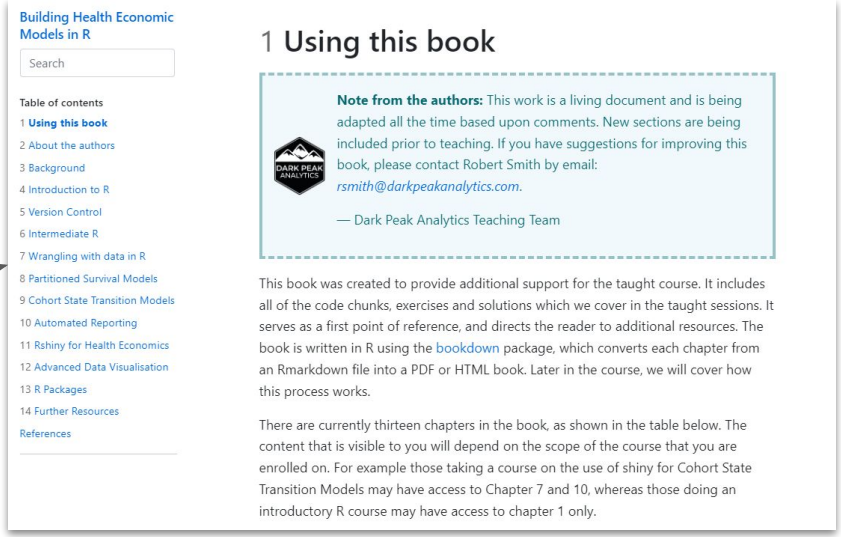
Case Study: NICE RCC Pilot



https://github.com/nice-digital/NICE-model-repo

Next steps

- Welcome contributions
  - Testing the software on your models
  - Suggesting improvements (see contribution page on GitHub)
  - Expansion of testing functionality
- Future development:
  - LLM Chatbot integration (using DPA teaching material to fine-tune).
  - Language selection (in progress)
- Open access publication imminent
- CRAN submission (Autumn ‘24)
- **Continued open-access development to maintain a collaborative tool**



Building Health Economic Models in R

Search

Table of contents

1 Using this book

Note from the authors: This work is a living document and is being adapted all the time based upon comments. New sections are being included prior to teaching. If you have suggestions for improving this book, please contact Robert Smith by email: *rsmith@darkpeakanalytics.com.*

— Dark Peak Analytics Teaching Team

This book was created to provide additional support for the taught course. It includes all of the code chunks, exercises and solutions which we cover in the taught sessions. It serves as a first point of reference, and directs the reader to additional resources. The book is written in R using the bookdown package, which converts each chapter from an Rmarkdown file into a PDF or HTML book. Later in the course, we will cover how this process works.

There are currently thirteen chapters in the book, as shown in the table below. The content that is visible to you will depend on the scope of the course that you are enrolled on. For example those taking a course on the use of shiny for Cohort State Transition Models may have access to Chapter 7 and 10, whereas those doing an introductory R course may have access to chapter 1 only.

https://github.com/dark-peak-analytics/assertHE

Pre-print available

Wellcome Open Research

Wellcome Open Research 2016 - DRAFT ARTICLE (PRE-SUBMISSION)

assertHE: an R package to improve quality assurance of HTA models

Robert A Smith*[1,2], Yevgeniy Samyshkin[3], Wael Mohammed[1,2], Felicity Lamrock[9], Tom Ward[3,4], Jack Smith[1], Alan Martin[3], Paul Schneider[1,2], Dawn Lee[5], Gianluca Baio[6], Howard Thom[7], Nathan Green[8], Marina Richardson[11], Mohamed El Alili[12,13], Xavier Pouwels[8], Calum Lewis[1,2], and Baris Deniz[3,10]

[1]Dark Peak Analytics, Sheffield, UK
[2]University of Sheffield, Sheffield, UK
[3]GSK, London, UK.
[4]University of Exeter Medical School, Exeter, EX1 2LU, UK
[5]PenTAG, University of Exeter Medical School, University of Exeter, St Luke's Campus, Exeter, EX1 2LU, UK
[6]Department of Statistical Science, University College London, London, UK
[7]Department of Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, England, UK
[8]Section of Health Technology and Services Research, Technical Medical Centre, Faculty of Behavioural, Management, and Social Sciences, University of Twente, Enschede, Overijssel, The Netherlands
[9]Mathematical Sciences Research Centre, Queen's University Belfast, UK
[10]Current affiliation: Alde Solutions LLC, 2810 N Church St PMB 824887, Wilmington, Delaware 19802-4447, US
[11]Institute for Clinical and Economic Review (ICER), Boston, MA.
[12]Department of Health Sciences, Faculty of Science, Vrije Universiteit Amsterdam, Amsterdam Public Health research Institute, Amsterdam, The Netherlands
[13]National Health Care Institute (ZIN), Diemen, The Netherlands

Corresponding author: Robert A Smith (rsmith@darkpeakanalytics.com).

Keywords

HTA, Health Economics, Unit Testing, Model Validation, Model Verification

Abstract

Background: Health economic models are increasingly used to inform decisions about the allocation of healthcare resources. Ensuring the robustness and reliability of these models is critical. Currently, quality assurance is conducted by both technical and non-technical experts assessing different components of the model manually. This is resource intensive. Understanding how the different components of the model fit together is time consuming, and testing every part of the model is sometimes not feasible in the time available. To aid in this, we have developed the `assertHE` R package.

Methods: The open source `assertHE` package provides testing functionality for those building and reviewing health economic models built in R programming language. It provides a series of checks which can be integrated into the model development workflow to reduce the probability of common errors. It also provides a suite of functions which allow users to better understand the network of R functions contained in a model, where they are defined, if (and where) they are tested, and provides a simple metric to quantify the extent to which they are tested.

Results: We applied the `assertHE` package to three open source health economic models built in R, showing how to include check functions within the model code and how to visualise the network of functions, see the test coverage, and obtain a Generative Pretrained Transformer Large Language Model (GPT-LLM) generated summary of any function in the code-base. We have worked with collaborators from industry, regulators and academia to develop the package to be applicable to the widest possible range of models, making adaptations to the source code based upon feedback.

Conclusions: `assertHE` offers an open-source toolkit for health economists building and reviewing models, promoting collaborative development and facilitating a more robust and efficient quality assurance process.

https://www.courses.darkpeakanalytics.com/products/digital_downloads/asserthe-quality-assurance

My experience

1. Much, much, better at summarising (dumbing down) than problem solving.
2. Will make things up, with confidence.
3. Volume matters - very good at small tasks. Hard to articulate nuance of bigger tasks.
4. Carefully written prompts are crucial, garbage in, garbage out.
5. Replacement for human Googling + Stack Overflow.

Potential

1. More efficient model builds.
2. Freeing up time for methods development.
3. Catching bugs & errors. If review > build then particularly useful for ERGs.
4. Making models more transparent to stakeholders?
5. Cross-pollination of ideas with other disciplines.

Challenges

1. Danger for open-source (does ‘no attribution’ destroy thought leadership?).

2. Will make things up, with confidence.

3. If simpler tasks (e.g. writing/reviewing specific functions) are done by genAI, how do people learn & develop?

4. How do we keep the human in the loop?! (Do we need to?)

Challenges

1. **Danger for open-source (does ‘no attribution’ destroy thought leadership?).**
2. Will make things up, with confidence.
3. If simpler tasks (e.g. writing/reviewing specific functions) are done by genAI, how do people learn & develop?
4. How do we keep the human in the loop?! (Do we need to?)

Challenges

1. Danger for open-source (does ‘no attribution’ destroy thought leadership?).

2. **Will make things up, with confidence.**

3. If simpler tasks (e.g. writing/reviewing specific functions) are done by genAI, how do people learn & develop?

4. How do we keep the human in the loop?! (Do we need to?)

Challenges

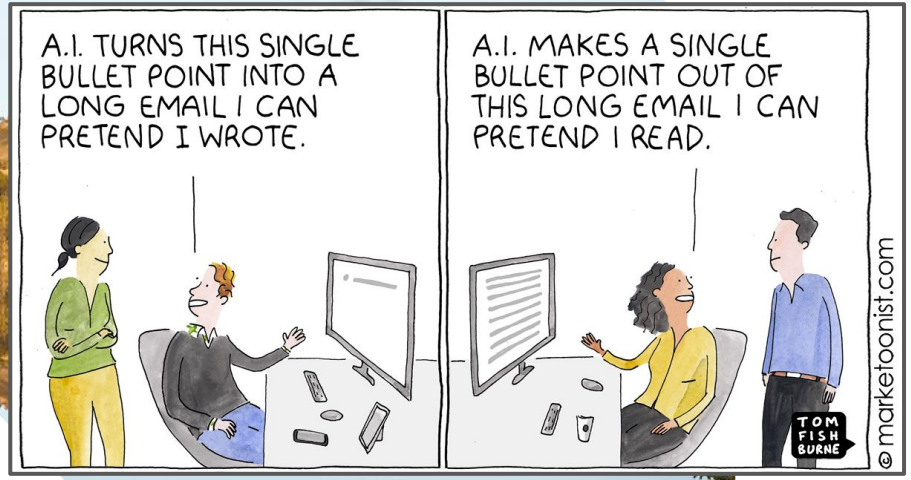
1. Danger for open-source (does ‘no attribution’ destroy thought leadership?).

2. Will make things up, with confidence.

3. **If simpler tasks (e.g. writing/reviewing specific functions) are done by genAI, how do people learn & develop?**

4. How do we keep the human in the loop?! (Do we need to?)

Challenges

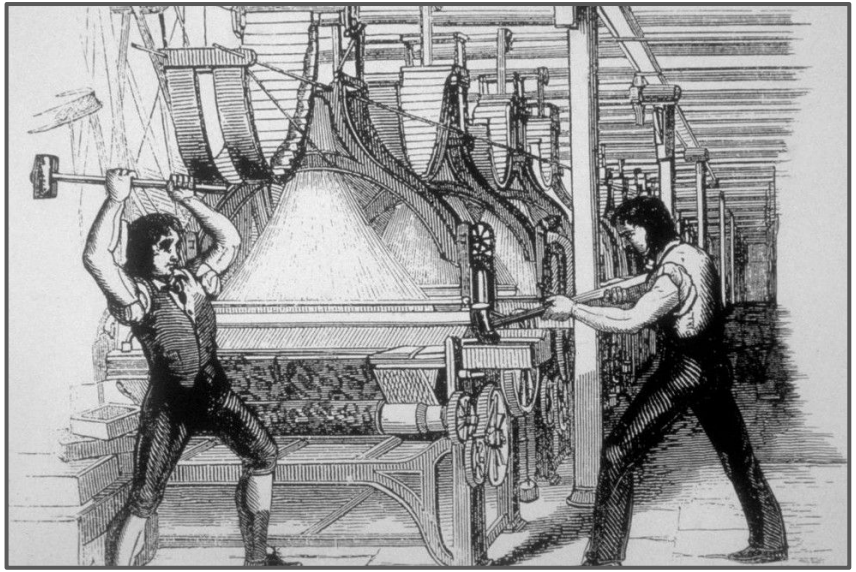


1. Danger for open-source (does ‘no attribution’ destroy thought leadership?).
2. Will make things up, with confidence.
3. If simpler tasks (e.g. writing/reviewing specific functions) are done by genAI, how do people learn & develop?
4. **How do we keep the human in the loop?! (Do we need to?)**

Challenges

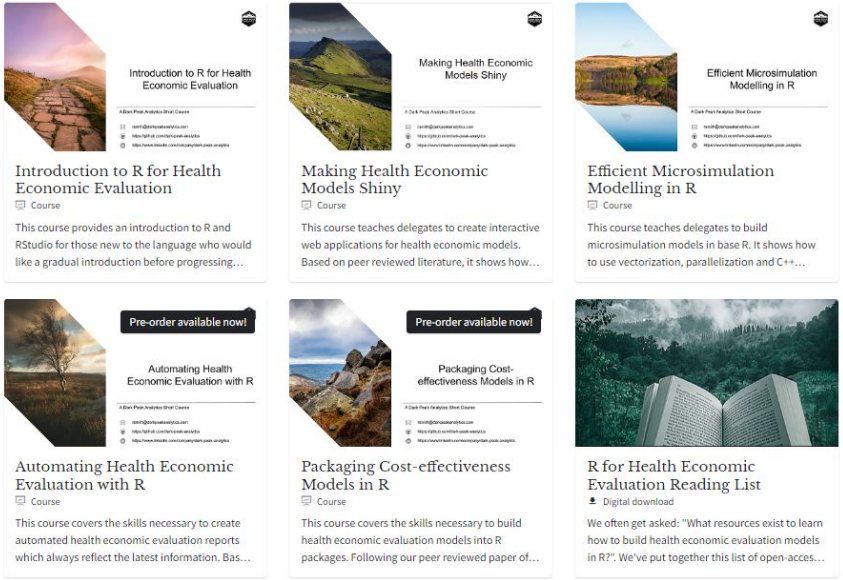
1. Danger for open-source (does ‘no attribution’ destroy thought leadership?).
2. Will make things up, with confidence.
3. If simpler tasks (e.g. writing/reviewing specific functions) are done by genAI, how do people learn & develop?
4. How do we keep the human in the loop?! (Do we need to?)
5. Will we all be out of work?

Challenges



5. **Will we all be out of work?**

Short Courses in R for Health Economic Evaluation



https://www.courses.darkpeakanalytics.com/